CLAIMS

1. An apparatus for detecting the presence of speech within an input audio signal, comprising:

a memory for storing a predetermined function which gives, for a given set of audio signal values, a probability density for parameters of a predetermined speech model which is assumed to have generated the set of audio signal values, the probability density defining, for a given set of model parameter values, the probability that the predetermined speech model has those parameter values, given that the speech model is assumed to have generated the set of audio signal values;

means for receiving a set of audio signal values representative of an input audio signal;

means for applying the set of received audio signal values to said stored function to give the probability density for said model parameters for the set of received audio signal values;

means for processing said function with said set of received audio signal values applied to obtain values of said parameters that are representative of said input audio signal; and

means for detecting the presence of speech using said obtained parameter values.

2. An apparatus according to claim 1, wherein said processing means comprises means for drawing samples from said probability density function and means for determining said values of said parameters that are representative of the speech from said drawn samples.

3. An apparatus according to claim 2, wherein said drawing means is operable to draw samples iteratively from said probability density function.

4. An apparatus according to claim 2, wherein said processing means comprises a Gibbs sampler.

5. An apparatus according to claim 2, wherein said processing means is operable to determine a histogram of said drawn samples and wherein said values of said parameters are determined from said histogram.

6. An apparatus according to claim 5, wherein said processing means is operable to determine said values of said parameters using a weighted sum of said drawn samples, and wherein the weighting is determined from said histogram.

7. An apparatus according to claim 1, wherein said

receiving means is operable to receive a sequence of sets of signal values representative of an input audio signal and wherein said applying means, processing means and detecting means are operable to perform their function with respect to each set of received audio signal values in order to determine whether or not each set of received signal values corresponds to speech.

8. An apparatus according to claim 7, wherein said processing means is operable to use the values of parameters obtained during the processing of a preceding set of signal values as initial estimates for the values of the corresponding parameters of a current set of signal values being processed.

9. An apparatus according to claim 7, wherein said sets of signal values in said sequence are non-overlapping.

10. An apparatus according to claim 1, wherein said speech model comprises an auto-regressive process model, wherein said parameters include auto-regressive model coefficients and wherein said detecting means is operable to compare the value of at least one of said auto-regressive model coefficients with a prestored threshold

value.

11. An apparatus according to claim 10, wherein said detecting means is operable to compare the values of a plurality of said auto-regressive model coefficients with a corresponding plurality of predetermined values.

12. An apparatus according to claim 1, wherein said processing means is operable to vary the number of parameters used to represent the speech within the audio signal values and wherein said detecting means is operable to compare the number of parameters used to represent speech within the audio signal values with a predetermined threshold value, in order to detect the presence of speech within said audio signal.

13. An apparatus according to claim 1, wherein received speech signal values are representative of a speech signal generated by a speech source as distorted by a transmission channel between the speech source and the receiving means; wherein said predetermined function includes a first part having first parameters which models said source and a second part having second parameters which models said channel; wherein said processing means is operable to obtain parameter values

of at least said first parameters; and wherein said detecting means is operable to detect the presence of speech within said input audio signal from the obtained values of said first parameters.

14. An apparatus according to claim 13, wherein said function is in terms of a set of raw speech signal values representative of speech generated by said source before being distorted by said transmission channel, wherein the apparatus further comprises second processing means for processing the received set of signal values with initial estimates of said first and second parameters, to generate an estimate of the raw speech signal values corresponding to the received set of audio signal values and wherein said applying means is operable to apply said estimated set of raw speech signal values to said function in addition to said set of received signal values.

15. An apparatus according to claim 14, wherein said second processing means comprises a simulation smoother.

16. An apparatus according to claim 14, wherein said second processing means comprises a Kalman filter.

17. An apparatus according to claim 13, wherein said second part is a moving average model and wherein said second parameters comprise moving average model coefficients.

18. An apparatus according to claim 1, further comprising means for evaluating said probability density function for the set of received audio signal values using one or more derived samples of parameter values for different numbers of parameter values, to determine respective probabilities that the predetermined speech model has those parameter values and wherein said processing means is operable to process at least some of said derived samples of parameter values and said evaluated probabilities to determine said values of said parameters that are representative of the audio speech signal.

19. A speech recognition system comprising:

means for receiving an input signal representative of an audio signal;

an apparatus according to claim 1 for detecting the presence of speech within the input signal; and

recognition processing means for performing a recognition processing of the portion of the input signal

corresponding to speech.

20. A speech processing system comprising:

means for receiving an input audio signal;

an apparatus according to claim 1 for detecting the presence of speech within the input audio signal; and

means for processing the portion of the input audio signal corresponding to speech.

21. A method of detecting the presence of speech within an input audio signal, comprising:

storing a predetermined function which gives, for a given set of audio signal values, a probability density for parameters of a predetermined speech model which is assumed to have generated the set of audio signal values, the probability density defining, for a given set of model parameter values, the probability that the predetermined speech model has those parameter values, given that the speech model is assumed to have generated the set of audio signal values;

receiving a set of audio signal values representative of an input audio signal at a receiver;

applying the set of received audio signal values to said stored function to give the probability density for said model parameters for the set of received audio

signal values;

processing said function with said set of received audio signal values applied to obtain values of said parameters that are representative of said input audio signal; and

detecting the presence of speech using said obtained parameter values.

22. A method according to claim 21, wherein said processing step comprises the steps of drawing samples from said probability density function and determining said values of said parameters that are representative of the speech from said drawn samples.

23. A method according to claim 22, wherein said drawing step draw samples iteratively from said probability density function.

24. A method according to claim 22, wherein said processing step uses a Gibbs sampler.

25. A method according to claim 22, wherein said processing step determines a histogram of said drawn samples and wherein said values of said parameters are determined from said histogram.

26. A method according to claim 25, wherein said processing step determines said values of said parameters using a weighted sum of said drawn samples, and wherein the weighting is determined from said histogram.

27. A method according to claim 21, wherein said receiving step receives a sequence of sets of signal values representative of an input audio signal and wherein said applying step, processing step and detecting step are performed on each set of received audio signal values in order to determine whether or not each set of received signal values corresponds to speech.

28. A method according to claim 27, wherein said processing step uses the values of parameters obtained during the processing of a preceding set of signal values as initial estimates for the values of the corresponding parameters of a current set of signal values being processed.

29. A method according to claim 27, wherein said sets of signal values in said sequence are non-overlapping.

30. A method according to claim 21, wherein said speech model comprises an auto-regressive process model, wherein

said parameters include auto-regressive model coefficients and wherein said detecting step compares the value of at least one of said auto-regressive model coefficients with a pre-stored threshold value.

31. A method according to claim 30, wherein said detecting step compares the values of a plurality of said auto-regressive model coefficients with a corresponding plurality of predetermined values.

32. A method according to claim 21, wherein said processing step varies the number of parameters used to represent the speech within the audio signal values and wherein said detecting step compares the number of parameters used to represent speech within the audio signal values with a predetermined threshold value, in order to detect the presence of speech within said audio signal.

33. A method according to claim 21, wherein received speech signal values are representative of a speech signal generated by a speech source as distorted by a transmission channel between the speech source and the receiver; wherein said predetermined function includes a first part having first parameters which models said

source and a second part having second parameters which models said channel; wherein said processing step obtains parameter values of at least said first parameters; and wherein said detecting step detects the presence of speech within said input audio signal from the obtained values of said first parameters.

34. A method according to claim 33, wherein said function is in terms of a set of raw speech signal values representative of speech generated by said source before being distorted by said transmission channel, wherein the apparatus further comprises a second processing step of processing the received set of signal values with initial estimates of said first and second parameters, to generate an estimate of the raw speech signal values corresponding to the received set of audio signal values and wherein said applying step applies said estimated set of raw speech signal values to said function in addition to said set of received signal values.

35. A method according to claim 34, wherein said second processing step uses a simulation smoother.

36. A method according to claim 34, wherein said second processing step uses a Kalman filter.

37. A method according to claim 33, wherein said second part is a moving average model and wherein said second parameters comprise moving average model coefficients.

38. A method according to claim 21, further comprising the step of evaluating said probability density function for the set of received audio signal values using one or more derived samples of parameter values for different numbers of parameter values, to determine respective probabilities that the predetermined speech model has those parameter values and wherein said processing step processes at least some of said derived samples of parameter values and said evaluated probabilities to determine said values of said parameters that are representative of the audio speech signal.

39. A speech recognition method comprising the steps of:

receiving an input signal representative of an audio signal;

a method according to claim 21 for detecting the presence of speech within the input signal; and

performing a recognition processing of the portion of the input signal corresponding to speech.

40. A speech processing method comprising the steps of:

receiving an input audio signal;

a method according to claim 21 for detecting the presence of speech within the input audio signal; and

processing the portion of the input audio signal corresponding to speech.

41. A computer readable medium storing computer executable process steps to cause a programmable computer apparatus to perform the method of claim 21.

42. Processor implementable process steps for causing a programmable computing device to perform the method according to claim 21.